Medi-Cal Management Information
System and Decision Support System (MIS/DSS)

Data Enhancement Functional Specifications for Patient Confidentiality

Phase 6

THE MEDSTAT® GROUP

December 4, 2000

Table of Contents

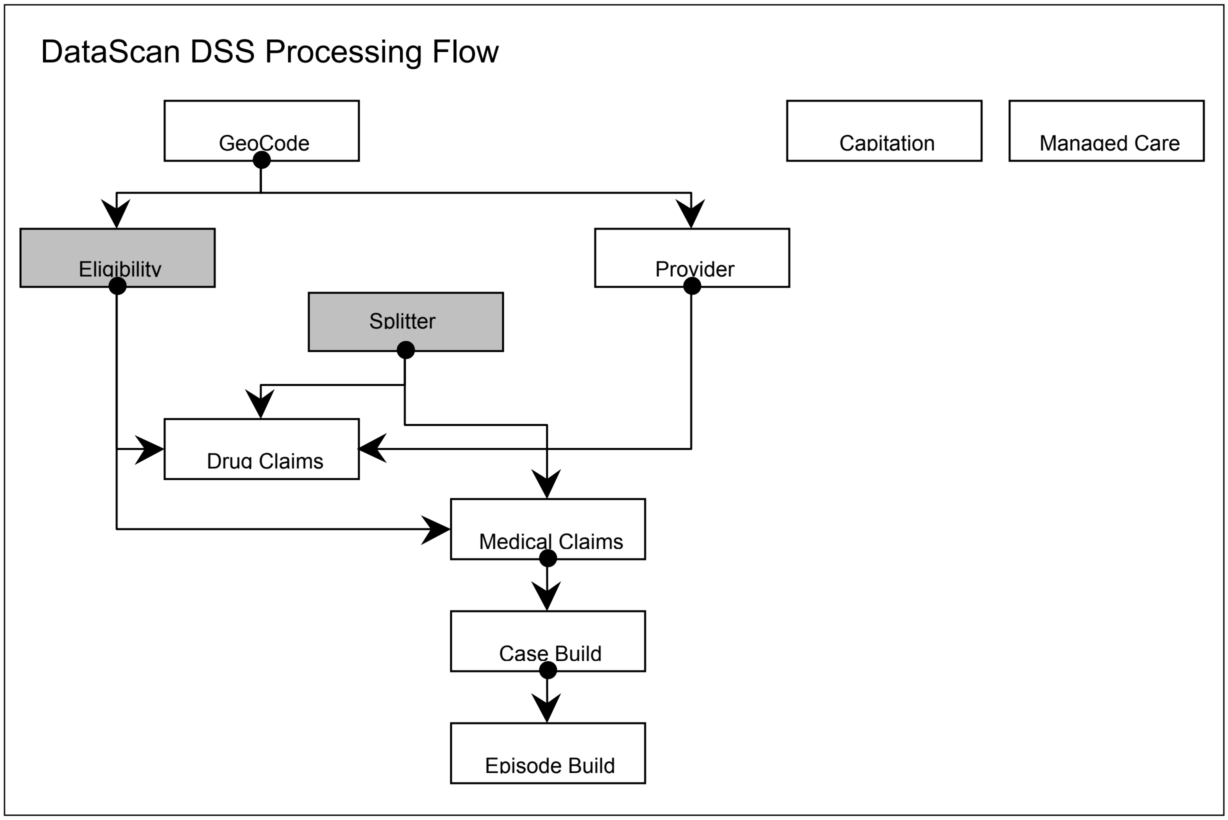
1. Overview

To protect patient confidentiality, all readily identifiable data in the MIS/DSS will either be encrypted or limited to secure users (identified by the Department of Health Services) on a need-to-know basis. This document describes the Encryption Process as well as the DB2 security views that will be created to restrict access.

The encryption process will be applied consistently to allow longitudinal analysis of individuals across the database. This process employs the use of the DHS Core Table which is used to store patient demographic information (e.g., name, mother's maiden name, DOB, etc). This table will be expanded by one column to carry both the original version of the CIN (new CINORIG field) as well as the encrypted version (current EMPID field). DataScan Study Group linking will enable secured users to select a group of individuals for investigation by the CINORIG field on the DHS Core Table and link to all the associated claims and eligibility information by the encrypted EMPID field. For users that are not at this level of security, all DSS table access will be via the encrypted EMPID field and no access will be granted to the DHS Core Table. Therefore the identity of a patient could not be readily determined by non-secure users. In addition, other key sensitive patient fields (e.g., name, SSNMEDS, etc) will be removed from table views for non-secure DSS users via DB2 column level security.

The CASENUM field, the last 10 bytes of the beneficiary ID, will also be encrypted. This field, along with the virtual fields created from multiple views of the CASENUM field are used to link families together. To retain the family associations, only the first 7 bytes of CASENUM will be encrypted. The original value of CASENUM (new field CASEORIG) as well as the new encrypted version (CASENUM) will be stored on the DHS Core Table. This will be two additional fields on the DHS Core Table. Section 5 of this document, Output Data, contains a complete, detailed list of fields that will be encrypted or hidden to non-secure DSS users.

The 'New Installation Considerations' section of this document will detail how this will be initially implemented in the current production database across all MIS/DSS products, as it will vary from the ongoing process.



2. Prerequisites / Pre-Conversion

GeoCode is a prerequisite to the Encryption Process. Refer to the Data Enhancement Functional Specifications for GeoCoding document for a more detailed description of the process.

3. Indexes

The new CINORIG field will be a secondary index on the DHS Core Table.

4. Input Data

Input to the Encryption process will be the output of the original EMPID convert validation in both the Eligibility and Splitter programs. The encryption of the first 7 bytes of the CASENUM field will only occur during the Eligibility Convert. This field is then tagged during the Drug and Claim Convert.

5. Output Data

The Eligibility Convert process outputs two intermediate files. 1.) The converted Eligibility file containing only the encrypted EMPID and CASENUM fields will be used to load the DataScan Eligibility Table as well as input to the Population build process. 2.) The file used to load the DHS Core Table will contain both the encrypted EMPID and CASENUM fields as well as the CINORIG and CASEORIG fields.

The Splitter Program will only output the encrypted EMPID field as input to the Drug and Claims convert. This encrypted EMPID will then be used in both the Case and Episodes Build processes.

The following is a complete list of fields in the DSS that will be affected by the process to protect patient confidentiality.

1. **BENEFRST** - On the DataScan Eligibility Table. This field will be hidden to non-secure DSS users.
2. **BENEINIT** - On the DataScan Eligibility Table. This field will be hidden to non-secure DSS users.
3. **BENELST** - On the DataScan Eligibility Table. This field will be hidden to non-secure DSS users.
4. **SSNMEDS -** On the DataScan Case, Dental, Drug, Eligibility, Episodes, IP and OP Service Tables. This field will be hidden to non-secure DSS users.

Note that the entire DHS Core Table will be hidden to non-secure DSS users.

5. **EMPID** - On the DataScan Case, Dental, Drug, Eligibility, Episodes, IP and OP Service (including Non-Dental OP) Tables. This entire field will be encrypted during the Splitter and Eligibility Convert programs.
6. **CINORIG** - On the DHS Core Table. Retains the original value of the Patient CIN.
7. **CASENUM -** On the DataScan Case, Dental, Drug, Eligibility, IP and OP Service Tables. The first 7 bytes of this 10 byte field will be encrypted during the Eligibility Convert process. This will ensure that the intended meaning behind the 8th, 9th and 10th bytes will remain intact. The virtual fields created from the CASENUM will be as follows:
   - **FBU** - view of the 8th byte (original value)
   - **FCASEFBU** - view of the first 8 bytes (first 7 bytes encrypted leaving the 8th byte with the original value)

- **FCASENBR** - view of the first 7 bytes (entire field encrypted)
- **PERSNBR** - view of the last 2 bytes (original values)

8. **CASEORIG** - On the DHS Core Table. Retains the original value of the CASENUM.

There will be no change to the current Panorama View and PMW extract programs as these are executed after the encryption will take place.

6. Reports

There are no additional reports produced by the Encryption Process.

7. Selection / Drop Criteria

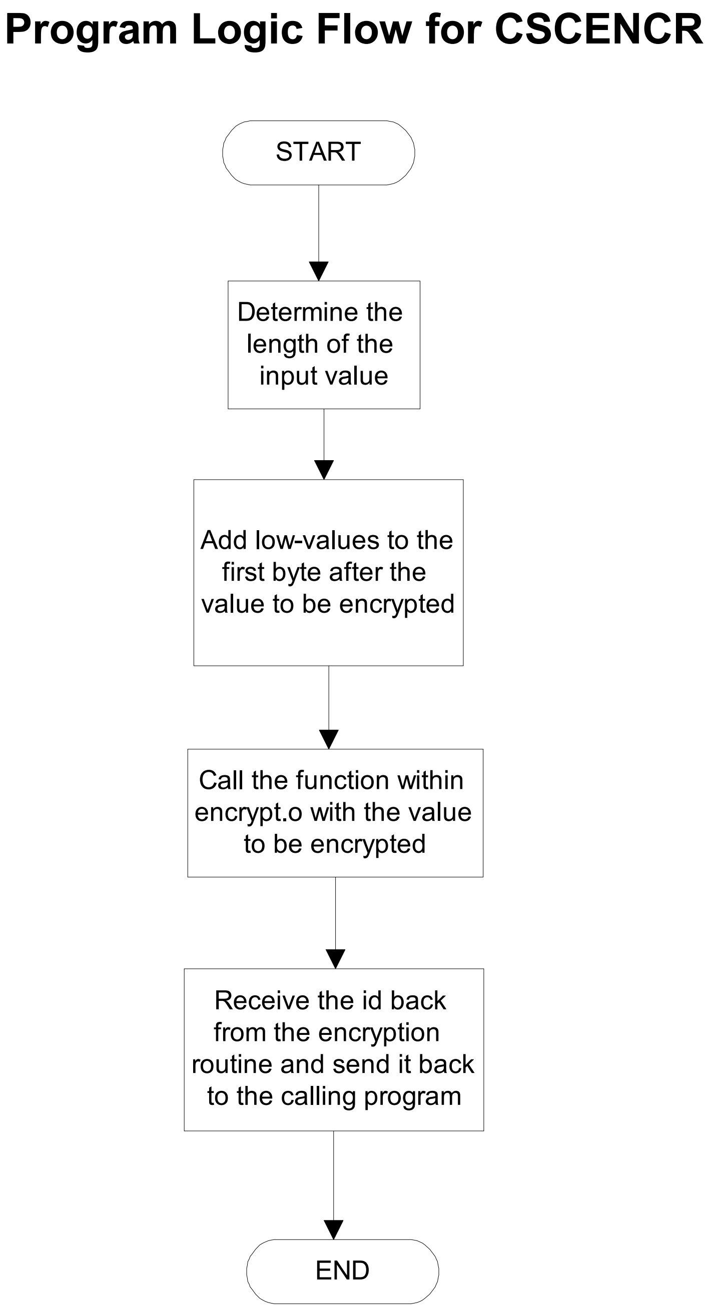
There is no Selection / Drop Criteria for the Encryption Process.

8. Process Flow / Data Enhancements

The Encryption routine will consist of 2 subroutines that take up to a 40 byte input field and produce up to a 40 byte encrypted output field.

In alignment with MEDSTAT's Privacy and Security Guidelines, the actual logic used to encrypt the input values can not be described in this document. A high-level process flow follows.

8.1 High-level Process Flow


Program Logic Flow for CSCENCR
START
Determine the length of the input value
Add low-values to the first byte after the value to be encrypted
Call the function within encrypt.o with the value to be encrypted
Receive the id back from the encryption routine and send it back to the calling program
END

The subroutine will take the input value and first determine its length. The input value must not contain spaces (Please note that this is a drop condition in the Splitter and Eligibility convert programs). The subroutine checks for the first byte that contains a space and that will determine the length of the field. If an EMPID has a space in the middle of the field, the encryption routine will treat that space as an end-of-field indicator.

The subroutine will then call the object module encrypt.o, which will then return the encrypted field and will be returned to the calling program as LK-OUTPUT-VALUE.

Note: The output value can only be moved to a character field and not a numeric because the encryption routine will return a combination of both characters and numeric data.

The Encryption Process will be initiated from within the Splitter and Eligibility Convert Programs. After the encryption routine has been called a separate check will take place to ensure that the number of characters output from the encryption routine is the same number that was input.

The DHS Core Table will act as the 'de-encryptor' for EMPID and CASENUM, allowing authorized users to cross-reference an encrypted ID with the original CIN. This will be the only place within the MIS/DSS that an original, unencrypted CIN or CASENUM can be used.

9. New Installation Considerations

9.1 DataScan and MyEureka

The implementation of the Encryption Process and the DB2 security views in the existing production database will vary from the ongoing monthly update process.

Each Production DataScan table will be unloaded to a 'working' area while the encryption takes place. A new field, CINORIG, will be created on the DHS Core Table which will contain a copy of the current EMPID value. Each of the DSS tables, including the HEDIS Service table used to archive claims for PMW, will then have an encryption routine executed in-place against their current EMPID field resulting in an encrypted ID replacement of the original EMPID field. The DHS Core Table will then have a cross reference map of the original CIN (CINORIG) and the encrypted CIN (EMPID). The net result of this effort would be that only secure users would have access rights to the DHS Core table that contains the actual real CIN.

The first 7 bytes of the CASENUM field on the Case, Dental, Drug, Eligibility, IP and OP Service Tables will also be encrypted during this process. The DHS Core Table will be expanded by two additional fields, the original CASENUM value (CASEORIG) and the encrypted version (CASENUM).

Once completed and a full backup has been done, a production outage will be required to switch the current production tables with the newly encrypted tables. No monthly updates will be able to occur during this time.

Both the Panorama View and Performance Measurement Workstation (PMW) databases will require special processing to encrypt the EMPID field.

9.2 Panorama View

Although Panorama View is an aggregated database, the internal tables contain EMPID. Once the encryption process is implemented, the DataScan extracts will contain encrypted EMPIDs. For Panorama View to be able to associate eligibility and claim records for the same beneficiary, all EMPIDs must be encrypted.

The implementation of the encryption process in Panorama View will not require production downtime. The Panorama View DB2 table will be unloaded to flat files. The encryption routine will run against the EMPID field in these flat files. The DB2 database will then be restored from the newly encrypted flat files. This can be done before the first Panorama View update using the DataScan extracts that contain the new encrypted EMPIDs.

9.3 Performance Measurement Workstation (PMW)

Patient level information is available within the PMW product for detailed analysis. To protect patient confidentiality, the input data used to build the PMW database will be encrypted and the database will be re-built.

The implementation of the encryption process in PMW will require small production downtime to move the encrypted database to the production server.

10. Update Processing Considerations

During a database update, the encryption of EMPID and CASENUM will occur in the Splitter and Eligibility Convert programs, respectively. All other update processes will remain the same.

11. Maps and Validation Tables

There are no maps or validation tables used in the Encryption Process.

12. Tagging

There are no fields tagged in the Encryption Process.

13. Summary of Document Changes

Date	Author	Phase	IRs	Description of Changes

Date	Author	Phase	IRs	Description of Changes
7/2/01	C.Swanson	6	1921	Corrected grammatical errors
1/9/01	C.Swanson	6	1921	Added clarification to Panorama View process.
1/02/01	C.Swanson	6	1921	Updated with the design change to retain the original value of CASENUM as well as the encrypted version on the DHS Core Table.
11/16/00	C.Swanson	6	1921	Initial creation